

Optimizing Constellations for Single-Subcarrier Intensity-Modulated Optical Systems

Johnny Karout, Erik Agrell, Krzysztof Szczerba, and Magnus Karlsson

Abstract—We optimize modulation formats for the additive white Gaussian noise channel with nonnegative input, also known as the intensity-modulated direct-detection channel, with and without confining them to a lattice structure. Our optimization criteria are the average electrical, average optical, and peak power. The nonnegative constraint on the input to the channel is translated into a conical constraint in signal space, and modulation formats are designed by sphere packing inside this cone. Some dense packings are found, which yield more power-efficient modulation formats than previously known. For example, at a spectral efficiency of 1.5 bit/s/Hz, the modulation format optimized for average electrical power has a 2.55 dB average electrical power gain over the best known format to achieve a symbol error rate of 10^{-6} . The corresponding gains for formats optimized for average and peak optical power are 1.35 and 1.72 dB, respectively. Using modulation formats optimized for peak power in average-power limited systems results in a smaller power penalty than when using formats optimized for average power in peak-power limited systems. We also evaluate the modulation formats in terms of their mutual information to predict their performance in the presence of capacity-achieving error-correcting codes, and finally show numerically and analytically that the optimal modulation formats for reliable transmission in the wideband regime have only one nonzero point.

Index Terms—Direct detection, fiber-optical communications, free-space optical communications, infrared communications, intensity modulation, lattice codes, mutual information, noncoherent communications, sphere packing.

I. INTRODUCTION

COHERENT optical systems, which give access to both the carrier's amplitude and phase to convey information, allow the design of higher-order modulation formats which offer a good trade-off between spectral and power efficiency. However, in systems where phase information is absent, designing such formats becomes challenging. Examples of such systems include phase-noise limited systems, and noncoherent systems where information is encoded onto the amplitude of the carrier and the envelope of the received signal is detected at the receiver. The latter is prevalent in optical communication systems where the overall cost and complexity is a critical constraint. This type of noncoherent systems is known as *intensity-modulated direct-detection* (IM/DD) systems and will be the focus of our work. In such systems, the information is encoded

onto the intensity of the optical carrier, and this intensity is, at all time instances, nonnegative. Applications using IM/DD are, for example, wireless optical communications [1]–[3] and short-haul fiber links used in, e.g., data centers [4], [5].

In the absence of optical amplification, an IM/DD system can be modeled as a conventional additive white Gaussian noise (AWGN) channel whose input is constrained to being nonnegative [1, Ch. 5], [2], [6]–[9]. Since the optical phase cannot be used to carry information, resorting to M -ary pulse amplitude modulation (M -PAM) is a natural low-complexity way of improving the spectral efficiency beyond that of the widespread on-off keying (OOK). However, this is different from the conventional PAM since no negative amplitudes can be used [1, Eq. (5.8)]. In [10], [11], an IM/DD link analysis using 4-PAM signaling was demonstrated. In [6], upper and lower bounds on the capacity of 2-, 4-, 8-, and 16-PAM were derived and in [12], the power efficiency of M -PAM was shown to be low. The M -ary pulse-position modulation (M -PPM) formats are known to be power-efficient; however, they suffer from poor spectral efficiency [1, Sec. 5.3.3], [2], [13].

Any nonnegative electrical waveform can be communicated successfully over an IM/DD link. This implies that if the information to be transmitted is modulated on an electrical subcarrier using any M -level modulation format, it can be transmitted on an IM/DD link after adding a direct current (DC) bias to ensure its nonnegativity, i.e., the subcarrier amplitude and phase which carry the information can be retrieved at the receiver. This concept is known as *subcarrier modulation* (SCM) and was described in the wireless infrared communications context [1, Ch. 5]. Therefore, the power efficiency compared to M -PAM can be improved since SCM allows the use of power-efficient higher-order modulation formats with IM/DD systems. In [14], the SCM concept was experimentally demonstrated, and in [15] and [16], a novel transmitter design for subcarrier quadrature phase-shift keying (QPSK) and 16-ary quadrature amplitude modulation (16-QAM) was presented. As for the conventional electrical channel, many subcarriers can be superimposed resulting in a frequency division-multiplexing (FDM) system, referred to as multiple-subcarrier modulation (MSM) in the wireless infrared context [1, p. 122], and orthogonal frequency-division multiplexing (OFDM) if the carriers are orthogonal [17]. Further, a subclass of OFDM known as discrete multitone (DMT), where the output of the inverse fast Fourier transform modulator is real instead of complex, was investigated in [18]. In [1, Sec. 5.3.2] and [19], MSM was shown to have poor power efficiency compared to single-subcarrier modulation. Specifically, Barry [1, Sec. 5.3.2] showed that the bandwidth

Manuscript received -; revised -. This work was supported by SSF under grant RE07-0026. The material in this paper was presented in part at the IEEE Global Communications Conference, Houston, TX, Dec. 2011.

J. Karout and E. Agrell are with the Department of Signals and Systems, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden e-mail: (johnny.karout, agrell@chalmers.se).

K. Szczerba and M. Karlsson are with the Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden e-mail: (krzysztof.szczerba, magnus.karlsson@chalmers.se).

of a multiple-subcarrier system is independent of the number of subcarriers, and that the average optical power penalty for multiple-subcarrier M -QAM in comparison with the single-subcarrier case is $5 \log_{10} N$ dB, where N is the number of subcarriers. In [20], DMT was shown to suffer considerably in peak-power limited systems. Kang and Hranilovic [21] and You and Kahn [22] proposed techniques to reduce the MSM power penalty.

For single-subcarrier modulation formats, one option is that the DC bias required to ensure the nonnegativity of the electrical waveform does not carry information [1, Ch. 5], [14], [16]. The second option is by allowing the DC bias to carry information, thus potentially improving the power efficiency. This was studied by varying the DC bias on a symbol-by-symbol basis in [23] and within the symbol interval in [22]. By guaranteeing nonnegativity, the investigation of *lattice codes* for IM/DD with AWGN became feasible [7], [24]. Lattice codes, which are finite sets of points selected out of an N -dimensional lattice, have been extensively used in the construction of higher-order modulation formats for AWGN channels with coherent detection [25]–[27]. In addition, techniques such as constellation shaping and nonequiprobable signaling have been used to further minimize the average power [26], [28]. In [7], a signal space model for optical IM/DD channels was presented, where average and peak optical power were considered as design constraints for constructing lattice-based modulation formats. Moreover, constellation shaping to reduce the average optical power were studied in [24] for the case where no amplification is used, and in [29] where optical amplifiers are used. An interesting question is if there exist new constellations that perform better than already known ones, with or without coding.

In this work, we address this question by optimizing single-subcarrier modulation formats for uncoded IM/DD systems, with and without confining them to a lattice structure. We choose a uniform probability distribution over constellation points as in [1]–[3], [7], [15], [16], [23], [30]–[32]. We propose a set of 4-, 8-, and 16-level single-subcarrier modulation formats which are optimized for average electrical, average optical, and peak power. These optimization criteria are all relevant, because the *average electrical* power is the standard power measure in digital and wireless communications [33, p. 40] and it helps in assessing the power consumption in optical communications [34], while the *average optical* power is an important figure of merit for skin- and eye-safety measures in wireless optical links [1, Ch. 5], [2], [7] and helps in quantifying the impact of shot noise in fiber-optical communications [35, p. 20]. In addition, the *peak* power, whether electrical or optical, is relevant for investigations of tolerance against the nonlinearities present in the system [36]. We then analyze the performance of the obtained modulation formats in terms of mutual information at different signal-to-noise ratios (SNR), in order to predict their performance in the presence of capacity-achieving error-correcting codes. Finally, we optimize modulation formats analytically in the wideband regime, i.e., at low SNR, while assuming uniform probability distributions and compare them with other formats.

The remainder of this paper is organized as follows. Sec-

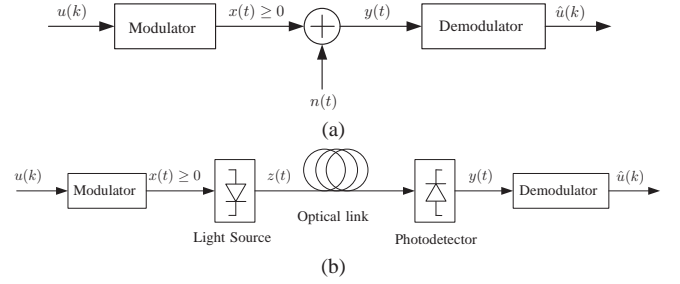


Fig. 1: (a) Baseband transceiver with constrained-input Gaussian channel. (b) Passband transceiver of IM/DD systems.

tion II presents the system model. Section III elaborates on the signal space model, the performance measures, and the single-subcarrier modulation family which is the focus of this work. Section IV explains the optimization criteria that were used and describes the obtained modulation formats. In Section V, we evaluate the performance of these modulation formats in the absence and presence of capacity-achieving error-correcting codes and we compare them with already known modulation formats. In addition, we analytically optimize modulation formats for the low-SNR regime. In Section VI, we summarize the main results and conclusions of this work.

II. SYSTEM MODEL

The system model under study is depicted in Fig. 1(a). It consists of a modulator which maps the symbol $u(k)$ at instant k to a waveform belonging to the signaling set $S = \{s_0(t), s_1(t), \dots, s_{M-1}(t)\}$, where T_s is the symbol period, M is the size of the signaling set, and $s_i(t) = 0$ for $t \notin [0, T_s]$ where $i = 0, 1, \dots, M-1$. For the time-disjoint signaling case, the generated waveform

$$x(t) = \sum_{k=-\infty}^{\infty} s_{u(k)}(t - kT_s), \quad (1)$$

where $u(k)$ is an ergodic process uniformly distributed over $\{0, 1, \dots, M-1\}$, is constrained to being real and nonnegative. The received signal can be written as

$$y(t) = x(t) + n(t), \quad (2)$$

where $n(t)$ is a zero-mean Gaussian process with double-sided power spectral density $N_0/2$. It should be noted that there exists no nonnegativity constraint on the signal $y(t)$. This is then followed by the demodulation of $y(t)$ which yields $\hat{u}(k)$, an estimate of $u(k)$. The demodulator is a correlator or matched filter receiver, which minimizes the symbol error rate at a given SNR [33, Sec. 4.1]. This model is different from the conventional AWGN channel by the fact that the input $x(t)$ is constrained to being nonnegative.

The baseband model in Fig. 1(a) has been extensively studied in the optical communications context, since it serves as a good model for IM/DD systems [1, Ch. 5], [2], [6]–[9], [37, Sec. 11.2.3]. The passband transceiver for IM/DD systems is depicted in Fig. 1(b). In such systems, the electrical nonnegative waveform $x(t)$ directly modulates a light source, such as a laser diode. Therefore, the information

is carried on the envelope of the passband signal $z(t) = \sqrt{2cx(t)} \cos(2\pi f_o t + \theta(t))$, i.e., the intensity of the optical field, where c represents the electro-optical conversion factor in watts per ampere (W/A) [35], [38], [39], f_o is the optical carrier frequency, and $\theta(t)$ is a random phase, uniformly distributed in $[0, 2\pi)$ and slowly varying with t . It then propagates through the optical medium depicted as an optical fiber in Fig. 1(b), which could be a free-space optical link in other applications. At the receiver, the photodetector detects the power of $z(t)$. Since the dominant channel impairment in optical IM/DD systems is the thermal noise resulting from the optical-to-electrical conversion [29], [40, p. 155], the received electrical signal can be written as

$$y(t) = rcx(t) + n(t), \quad (3)$$

where r is the responsivity of the opto-electrical converter in A/W. Without loss of generality, we set $rc = 1$, which yields (2).

There exists another IM/DD model which is relevant when the dominating noise comes from optical amplifiers, and not the receiver [37, Sec. 11.2], [41], [42]. The noise in that model has a noncentral χ^2 -distribution and not a Gaussian distribution as in this work.

III. SIGNAL SPACE MODEL

By defining a set of orthonormal basis functions $\phi_k(t)$ for $k = 1, 2, \dots, N$ and $N \leq M$ as in [7], each of the signals in S can be represented as

$$s_i(t) = \sum_{k=1}^N s_{i,k} \phi_k(t) \quad (4)$$

for $i = 0, \dots, M-1$, where $\mathbf{s}_i = (s_{i,1}, s_{i,2}, \dots, s_{i,N})$ is the vector representation of $s_i(t)$ with respect to the aforementioned basis functions. Therefore, the constellation representing the signaling set S can be written as $\Omega = \{\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{M-1}\}$. With this representation, the continuous-time channel models in (2) and (3) can be represented by the discrete-time vector model

$$\mathbf{y}(k) = \mathbf{x}(k) + \mathbf{n}(k), \quad (5)$$

where, at instant k , $\mathbf{x}(k) \in \Omega$ is the transmitted vector and $\mathbf{n}(k)$ is a Gaussian random vector with independent elements, zero mean, and variance $N_0/2$ per dimension. Since $\mathbf{x}(k)$ and $\mathbf{y}(k)$ are both stationary processes, the argument k will be dropped from now on. To satisfy the nonnegativity constraint of the channel, the basis function $\phi_1(t)$ is set as in [6], [7] to

$$\phi_1(t) = \sqrt{\frac{1}{T_s}} \text{rect}\left(\frac{t}{T_s}\right), \quad (6)$$

where

$$\text{rect}(t) = \begin{cases} 1, & \text{if } 0 \leq t \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

This basis function represents the DC bias. Thus, $s_{i,1}$ is chosen for each $i = 0, \dots, M-1$ such that

$$\min_t s_i(t) \geq 0,$$

which guarantees the nonnegativity of $x(t)$ in (1). The admissible region Υ containing the set of all signal vectors satisfying the nonnegativity constraint can be represented as [7, Eq. (10)]

$$\Upsilon = \{\mathbf{w} \in \mathbb{R}^N : \min_{t \in [0, T_s]} \sum_{k=1}^N w_k \phi_k(t) \geq 0\}, \quad (7)$$

where $\mathbf{w} = (w_1, w_2, \dots, w_N)$. Therefore, the constellation Ω is a finite subset of Υ . The admissible region Υ for IM/DD systems has been shown in [7, Th. 1] to be the convex hull of a generalized N -dimensional cone with vertex at the origin and opening in the dimension spanned by $\phi_1(t)$.

A. Performance Measures

Unlike the conventional electrical AWGN channel where the two standard power performance measures are the average and peak electrical power, three important performance measures for IM/DD channels can be extracted from the baseband and passband models in Fig. 1. The first entity is the average electrical power defined as

$$\bar{P}_e = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x^2(t) dt,$$

which for any basis functions can be simplified to

$$\bar{P}_e = \frac{E_s}{T_s} = \frac{1}{T_s} \mathbb{E}[\|s_I\|^2], \quad (8)$$

where E_s is the average energy of the constellation, $\mathbb{E}[\cdot]$ is the expected value, and I is a random variable uniformly distributed over $\{0, 1, \dots, M-1\}$. This entity is an important figure of merit for assessing the performance of digital and wireless communication systems [33, p. 40]. Therefore, it is relevant for IM/DD systems for compatibility with classical methods and results [42], [43]. In addition, it helps in quantifying the impact of relative intensity noise (RIN) in fiber-optical links [35], and in assessing the power consumption of optical systems [34]. In [30], \bar{P}_e was used as a performance measure for comparing different intensity modulation formats.

The second measure is the average optical power \bar{P}_o , which has been studied in [1], [2], [6]–[8] for the wireless optical channel. Limitations are set on \bar{P}_o for skin- and eye-safety standards to be met. In fiber-optic communications, this entity is used to quantify the impact of shot noise on the performance [35, p. 20]. It is defined as

$$\bar{P}_o = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T z^2(t) dt = \lim_{T \rightarrow \infty} \frac{c}{2T} \int_{-T}^T x(t) dt.$$

This measure depends solely on the DC bias required to make the signals nonnegative and can be represented in terms of the symbol period and constellation geometry as [6], [7]

$$\bar{P}_o = \frac{c}{\sqrt{T_s}} \mathbb{E}[s_{I,1}], \quad (9)$$

regardless of $\phi_2(t), \dots, \phi_N(t)$.

The third measure is the peak optical power defined as

$$\begin{aligned} \hat{P}_o &= \max_t \frac{z^2(t)}{2} \\ &= c \max_t x(t). \end{aligned} \quad (10)$$

It is relevant for investigations of tolerance against the non-linear behavior of transmitting and receiving hardware in communication systems [7], [36], [39] and has been studied in [7], [8], [30]. The peak electrical power \hat{P}_e is directly related to \hat{P}_o by

$$\hat{P}_e = \left(\frac{\hat{P}_o}{c} \right)^2,$$

and will not be further considered in this paper, since a constellation optimized for \hat{P}_o will automatically be optimized for \hat{P}_e too. A general form for \hat{P}_o as a function of Ω , as in (8) and (9) for \bar{P}_e and \bar{P}_o , does not exist, since \hat{P}_o depends on the exact choice of basis functions. A special case will be studied in Sec. III-B.

To assess the performance of the different modulation formats in the presence of capacity-achieving error-correcting codes, we consider the mutual information [44, Sec. 2.4]

$$I(\mathbf{x}; \mathbf{y}) = H(\mathbf{x}) - H(\mathbf{x}|\mathbf{y}) \quad (11)$$

as a performance measure. The terms $H(\mathbf{x})$ and $H(\mathbf{x}|\mathbf{y})$ are the entropy of \mathbf{x} and the conditional entropy of \mathbf{x} given the received vector \mathbf{y} , averaged over both \mathbf{x} and \mathbf{y} . The channel capacity of a discrete memoryless channel is [44, Eq. (7.1)]

$$C = \max_{p(\mathbf{x})} I(\mathbf{x}; \mathbf{y}), \quad (12)$$

where the maximum is taken over all possible input distributions $p(\mathbf{x})$. For a fixed constellation and distribution, the mutual information gives a lower bound on the channel capacity. In our work, like in the works of many other authors, we choose a uniform distribution over the constellation points.

We define $R_s = 1/T_s$ as the symbol rate in symbols per second, $R_b = R_s R$ as the bit rate in bits per second, and $E_b = E_s/R$ as the average energy per bit. Furthermore, in order to have a fair comparison of the bit rates that can be achieved by the different modulation formats in a fixed bandwidth, the spectral efficiency defined as

$$\eta = \frac{R_b}{W} \text{ [bit/s/Hz]}$$

should be taken into account, where W is the baseband bandwidth defined as the first null in the spectrum of $x(t)$. In this paper, we are interested in two extreme cases: the uncoded system, for which $R = \log_2 M$, and the system with optimal coding, for which $R = I(\mathbf{x}; \mathbf{y})$.

B. Single-Subcarrier Modulation Formats

For in-phase and quadrature phase (I/Q) modulation formats to be used on intensity modulated channels, a DC bias is required in order for $x(t)$ to be nonnegative. This could be translated geometrically by having a three-dimensional (3d) Euclidean space spanned by the orthonormal basis functions $\phi_1(t)$ defined in (6) and

$$\phi_2(t) = \sqrt{\frac{2}{T_s}} \cos(2\pi f t) \text{ rect}\left(\frac{t}{T_s}\right), \quad (13)$$

$$\phi_3(t) = \sqrt{\frac{2}{T_s}} \sin(2\pi f t) \text{ rect}\left(\frac{t}{T_s}\right), \quad (14)$$

which are the basis functions of conventional I/Q modulation formats such as M -PSK and M -QAM, where f is the electrical subcarrier frequency [1], [7]. As in [1, pp. 115–116] and [7], we use $f = 1/T_s$, which is the minimum value for which $\phi_1(t)$, $\phi_2(t)$, and $\phi_3(t)$ are orthonormal. In [7], IM/DD modulation formats based on these three basis functions are referred to as raised-QAM, and in [14] as single cycle SCM. At the same symbol rate, modulation formats such as OOK and M -PAM have $W = R_s$, whereas the modulation formats belonging to the single-subcarrier family occupy $W = 2R_s$; this is due to the intermediate step of modulating the information onto an electrical subcarrier before modulating the optical carrier [1, Ch. 5], [30].

We now describe explicitly the admissible region Υ for single-subcarrier modulation formats [3, Fig. 4.2], [7].

Theorem 1: For the specific set of basis functions $\phi_1(t)$, $\phi_2(t)$, and $\phi_3(t)$ defined in (6), (13), and (14), the admissible region Υ is a three-dimensional (3d) cone with vertex at the origin, apex angle of $\cos^{-1}(1/3) = 70.528^\circ$, and opening in the dimension spanned by $\phi_1(t)$.

Proof: The admissible region in (7) can be written for single-subcarrier modulation formats as

$$\Upsilon = \{ \mathbf{w} \in \mathbb{R}^3 : \min_{t \in [0, T_s)} \sum_{k=1}^3 w_k \phi_k(t) \geq 0 \}, \quad (15)$$

where

$$\begin{aligned} & \min_{t \in [0, T_s)} \sum_{k=1}^3 w_k \phi_k(t) \\ &= \frac{1}{\sqrt{T_s}} \min_{t \in [0, T_s)} \{ w_1 \\ & \quad + \sqrt{2(w_2^2 + w_3^2)} \cos(2\pi f t - \theta) \} \\ &= \frac{1}{\sqrt{T_s}} (w_1 - \sqrt{2(w_2^2 + w_3^2)}), \end{aligned} \quad (16)$$

where $\theta = \arg(w_2 + jw_3)$. Therefore, substituting (16) in (15) yields

$$\Upsilon = \{ \mathbf{w} \in \mathbb{R}^3 : w_1 \geq \sqrt{2(w_2^2 + w_3^2)} \}, \quad (17)$$

which is a 3d-cone with apex angle of $\cos^{-1}(1/3) = 70.528^\circ$ pointing in the dimension spanned by $\phi_1(t)$, with vertex at the origin. ■

The average electrical and optical power were given in (8) and (9) as functions of the constellation Ω ; however, the peak optical power defined in (10) could for these basis functions be expressed in terms of the constellation geometry too [3, Fig. 4.3].

Theorem 2: The peak optical power for the single-subcarrier modulation formats with the above defined basis functions can be expressed as

$$\hat{P}_o = \frac{c}{\sqrt{T_s}} \max_i \left\{ s_{i,1} + \sqrt{2(s_{i,2}^2 + s_{i,3}^2)} \right\}. \quad (18)$$

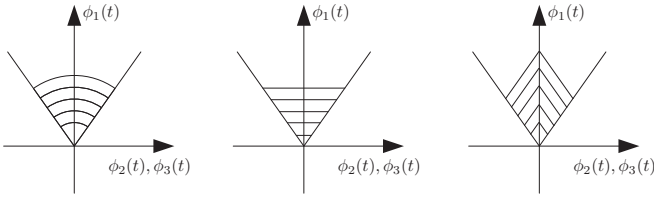


Fig. 2: (left to right): Contours of equal \bar{P}_e , \bar{P}_o , and \hat{P}_o .

Proof: From (10), \hat{P}_o can be written as

$$\begin{aligned}\hat{P}_o &= c \max_{i,t} s_i(t) \\ &= c \max_{i,t} \sum_{k=1}^3 s_{i,k} \phi_k(t) \\ &= \frac{c}{\sqrt{T_s}} \max_i \left\{ s_{i,1} + \sqrt{2(s_{i,2}^2 + s_{i,3}^2)} \right\}.\end{aligned}$$

Alternatively, the theorem can be proved using [7, Th. 2]. ■

IV. CONSTELLATION OPTIMIZATION

To design power-efficient constellations, the admissible region in (17) has to be taken into account. As done before for the conventional AWGN channel [45]–[49], our approach of finding the best constellations can be formulated as a sphere-packing problem with the objective of minimizing a cost function depending on the constraints that might be present in the system model shown in Fig. 1. Thus, the optimization problem, for given constants M and d_{\min} , can be written as

$$\text{Minimize} \quad \xi(\Omega) \quad (19)$$

$$\text{Subject to} \quad |\Omega| = M \quad (20)$$

$$\Omega \subset \Upsilon \quad (21)$$

$$d(\Omega) = d_{\min}, \quad (22)$$

where

$$d(\Omega) = \min_{\substack{\mathbf{s}_i, \mathbf{s}_j \in \Omega \\ i \neq j}} \|\mathbf{s}_i - \mathbf{s}_j\|.$$

Choosing the objective function as $\xi(\Omega) = \mathbb{E}[\|s_I\|^2]$ results in $\Omega = \mathcal{C}_{\bar{P}_e, M}$, i.e., a constellation optimized for average electrical power, and $\xi(\Omega) = \mathbb{E}[s_{I,1}]$ results in $\Omega = \mathcal{C}_{\bar{P}_o, M}$, i.e., a constellation optimized for average optical power. Finally,

$$\xi(\Omega) = \max_i \left\{ s_{i,1} + \sqrt{2(s_{i,2}^2 + s_{i,3}^2)} \right\}$$

yields $\Omega = \mathcal{C}_{\hat{P}_o, M}$, a constellation optimized for peak optical power. Fig. 2 depicts a two-dimensional contour plot of the three objective functions together with the admissible region Υ . The constraint in (21) guarantees that the signals belong to the admissible region Υ , therefore satisfying the nonnegativity criterion of the channel. The minimum distance d_{\min} in (22) serves as a good measure of error probability performance in the presence of AWGN at high SNR. Although this optimization problem is well formulated mathematically, it is difficult to obtain an analytical solution. Therefore, we resorted to *numerical optimization* techniques as in [45]–[49] to find

the best constellations. One drawback of such constellations is often the lack of geometric regularity, which increases the modulator and demodulator complexity. The optimization problem is nonconvex; therefore, a local solution does not imply that it is globally optimal.

A special case of this optimization problem, which might not guarantee the optimal solution, is to confine the possible constellations to have a regular structure such as that of a lattice, denoted by Λ . In this case, the above optimization problem can be reformulated by replacing (21) with $\Omega \subset \Upsilon \cap \Lambda$, and dropping (22) since it is directly inferred by $\Omega \subset \Lambda$. This was done in [7] for the cubic and Leech lattices. In order to compare with the best nonlattice constellations of relatively small sizes M , we use in this work the face-centered cubic lattice (A_3), which provides the densest packing for the 3d-Euclidean space [27, p. xvi]. The obtained lattice-based constellations optimized for average electrical, average optical, and peak optical power are denoted $\mathcal{L}_{\bar{P}_e, M}$, $\mathcal{L}_{\bar{P}_o, M}$, and $\mathcal{L}_{\hat{P}_o, M}$, respectively.

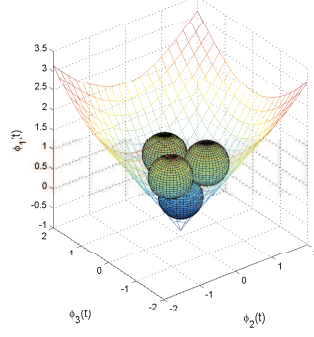
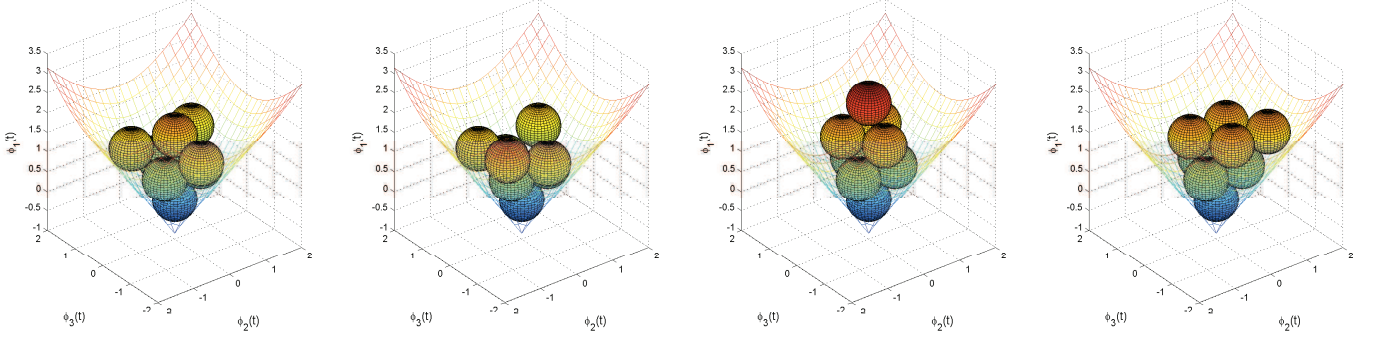
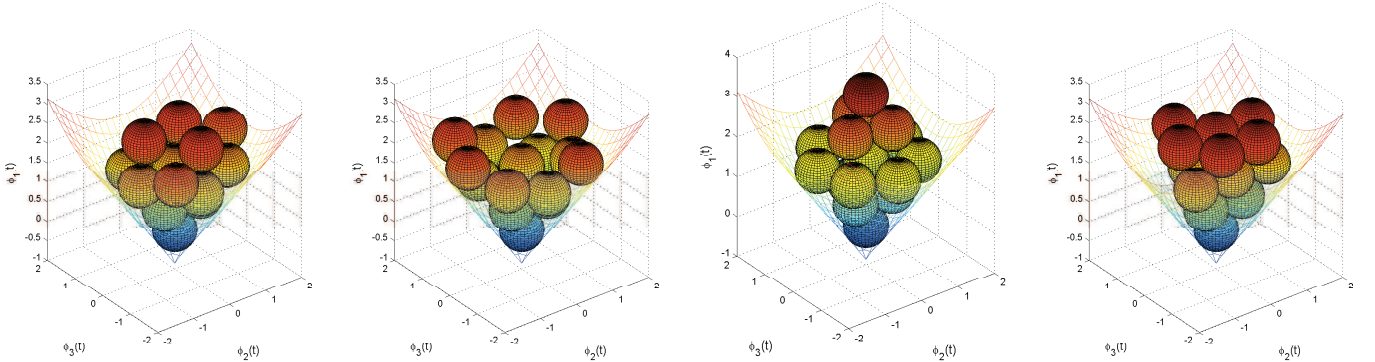
A. Optimized Constellations

In Figs. 3–5, the results of the numerical optimizations are illustrated for $M = 4, 8$, and 16, for unconstrained sphere packings (\mathcal{C}) and lattice codes (\mathcal{L}), and for the three power measures. Their coordinates are included in App. A. It should be noted that rotations of all constellations about $\phi_1(t)$ do not change the power requirements. As we shall see in Sec V, the obtained constellation outperform previously known formats.

Conjecture 3: All constellations in App. A are optimal solutions of (19)–(22).

1) *4-level Constellations:* The same 4-level constellation provides the lowest \bar{P}_e , \bar{P}_o , and \hat{P}_o while satisfying the optimization constraints. The geometry of this constellation is a regular tetrahedron where all the spheres, or the constellation points lying at the vertices of this regular tetrahedron, are equidistant from each other. This constellation is also the result of the optimization constrained to $\Omega \subset \Upsilon \cap A_3$, where the apex of the cone coincides with a point in the A_3 lattice and the lattice is oriented such that two lattice basis vectors lie in the plane spanned by $\phi_2(t)$ and $\phi_3(t)$. Since the obtained constellation is optimized for \bar{P}_e , \bar{P}_o , and \hat{P}_o , we will refer to it as \mathcal{C}_4 or \mathcal{L}_4 .

It is a remarkable fact that the vertex angle of the tetrahedron, defined as the apex angle of the circumscribed cone, is exactly $\cos^{-1}(1/3)$, which is equal to the apex angle of the admissible region Υ . Thus, \mathcal{C}_4 fits Υ snugly, in the sense that all constellation points are equidistant from each other and lie on the boundary of Υ . For constellation points regarded as unit-diameter spheres, \mathcal{C}_4 can be illustrated as four spheres touching each other and the boundary of a larger cone as shown in Fig. 3. This, as we shall see in the next section, makes the modulation format very power-efficient. This modulation format consists of a zero-level signal and a biased ternary PSK constellation [50], [51]. In prior work [30], the \mathcal{C}_4 format was introduced where it was called on-off phase-shift keying (OOPSK), and in [31], it was demonstrated experimentally. Other hybrids between amplitude-shift keying

Fig. 3: $\mathcal{C}_4 = \mathcal{L}_4$.Fig. 4: (left to right): $\mathcal{C}_{\bar{P}_e,8}$, $\mathcal{C}_{\bar{P}_o,8}$, $\mathcal{C}_{\hat{P}_o,8} = \mathcal{L}_{\bar{P}_o,8}$, $\mathcal{L}_{\bar{P}_e,8} = \mathcal{L}_{\bar{P}_o,8}$.Fig. 5: (left to right): $\mathcal{C}_{\bar{P}_e,16}$, $\mathcal{C}_{\bar{P}_o,16}$, $\mathcal{C}_{\hat{P}_o,16}$, \mathcal{L}_{16} .

and PSK have been studied in [32] and [52]; however, such modulation formats do not satisfy the nonnegativity constraint of IM/DD channels.

2) *8-level Constellations*: The highly symmetric and compact constellation $\mathcal{C}_{\bar{P}_e,8}$ consists of four central spheres arranged in a tetrahedron and four additional spheres, each touching three spheres in the central tetrahedron. Surprisingly, seven of the eight spheres touch the conical boundary of Υ . This modulation format is a hybrid between 2-PAM and two ternary PSK constellations, which are DC-biased differently. The constellation $\mathcal{C}_{\bar{P}_o,8}$ is the same as $\mathcal{C}_{\bar{P}_e,8}$ but with the top central sphere moved to the boundary of the admissible region. The constellation optimized for peak optical power, $\mathcal{C}_{\hat{P}_o,8}$, consists of two tetrahedra lying on top of each other, where one is reflected and rotated $\pi/3$ about $\phi_1(t)$.

On the other hand, when confining the set of points to a lattice structure, the resulting constellations which provide the lowest \bar{P}_e and \bar{P}_o are the same, $\mathcal{L}_{\bar{P}_e,8} = \mathcal{L}_{\bar{P}_o,8}$. However, the lattice-based constellation which is optimized for \hat{P}_o , $\mathcal{L}_{\hat{P}_o,8}$, is the same as $\mathcal{C}_{\hat{P}_o,8}$.

3) *16-level Constellations*: The constellations $\mathcal{C}_{\bar{P}_e,16}$, $\mathcal{C}_{\bar{P}_o,16}$, and $\mathcal{C}_{\hat{P}_o,16}$ are not lattice codes; however, the shape obtained could be well justified by the contour plots shown in Fig. 2. The $\mathcal{C}_{\bar{P}_o,16}$ constellation contains the constellations $\mathcal{C}_{\bar{P}_o,8}$ and $\mathcal{C}_{\bar{P}_e,8}$, whereas $\mathcal{C}_{\hat{P}_o,16}$ consists of two tetrahedra of which one is reflected and eight spheres lying in between, almost at the same level. The presence of two tetrahedra in the constellations optimized for peak optical power, whether 8- or 16-levels, is due to the fact that these constellations are bounded from above by another cone with the same apex angle

as that of Υ . This can be inferred from the expression of \hat{P}_o in (18) [7, Th. 2].

However, when only lattice-based structures are considered, the constellations providing the lowest \bar{P}_e , \bar{P}_o , \hat{P}_o are the same, i.e., $\mathcal{L}_{16} = \mathcal{L}_{\bar{P}_e,16} = \mathcal{L}_{\bar{P}_o,16} = \mathcal{L}_{\hat{P}_o,16}$. From Figs. 4–5, it can be noticed that \mathcal{L}_{16} contains both $\mathcal{C}_{\bar{P}_o,8}$ and $\mathcal{L}_{\bar{P}_e,8}$, and that the \mathcal{C}_4 constellation is included in all the obtained constellations.

Conjecture 4: For single-subcarrier IM/DD systems, the \mathcal{C}_4 constellation is included in all optimal constellations with $M \geq 4$.

B. Previously Known Constellations

Our investigation encompasses some previously best known formats, which are presented after being normalized to unit d_{\min} . Readers are referred to [53] for the performance of classical formats such as M -PSK and M -QAM over IM/DD channels. At spectral efficiency $\eta = 1$ bit/s/Hz (where $R = \log_2 M$), OOK is defined as $\{(0), (1)\}$ in terms of $\phi_1(t)$.

At spectral efficiency $\eta = 1.5$ bit/s/Hz, a star-shaped 8-QAM [32] denoted as 8-QAM, in which the DC bias is allowed to vary from symbol to symbol, is defined as $\{(1, \pm 1/2, \pm 1/2), ((1 + \sqrt{3})/\sqrt{2}, 0, \pm(1 + \sqrt{3})/2), ((1 + \sqrt{3})/\sqrt{2}, \pm(1 + \sqrt{3})/2, 0)\}$. At spectral efficiency $\eta = 2$ bit/s/Hz, nonnegative 4-PAM is defined as $\{(0), (1), (2), (3)\}$ in terms of $\phi_1(t)$, and a version of 16-QAM denoted as 16-QAM where the DC bias varies from symbol to symbol, is defined as $\{(1, \pm 1/2, \pm 1/2), (\sqrt{5}, \pm 1/2, \pm 3/2), (\sqrt{5}, \pm 3/2, \pm 1/2), (3, \pm 3/2, \pm 3/2)\}$.

V. PERFORMANCE ANALYSIS

In this section, we assess the performance of all the modulation formats considered in this work in terms of their symbol error rate (SER) performance, asymptotic power gain versus OOK, which was used as a benchmark in [1], [2], and spectral efficiency. Sections V-A–V-B consider uncoded transmission ($R = \log_2 M$) and Sections V-C–V-D consider coded transmission ($R = I(\mathbf{x}; \mathbf{y})$).

A. Symbol Error Rate

For the 4-level modulation \mathcal{C}_4 presented in Fig. 3, deriving the exact theoretical SER is not straight-forward, due to the irregularity of the Voronoi regions. However, it has the same structure as the simplex signal set in [33, Sec. 4.1], although it is DC-biased to be used in IM/DD systems. The exact SER of an M -ary simplex signal set is [33, Eq. (4.116)]

$$P_s = 1 - \int_{-\infty}^{\infty} \left[1 - Q \left(u + \sqrt{\frac{2E_{s,\text{simplex}}}{N_0} \frac{M}{M-1}} \right) \right]^{M-1} \cdot \frac{e^{-u^2/2}}{\sqrt{2\pi}} du, \quad (23)$$

where $E_{s,\text{simplex}}$ is the average symbol energy of the zero-mean simplex constellation and $Q(x) = 1/\sqrt{2\pi} \int_x^{\infty} \exp(-u^2/2) du$ is the Gaussian Q-function. For $M = 4$, the energy of the

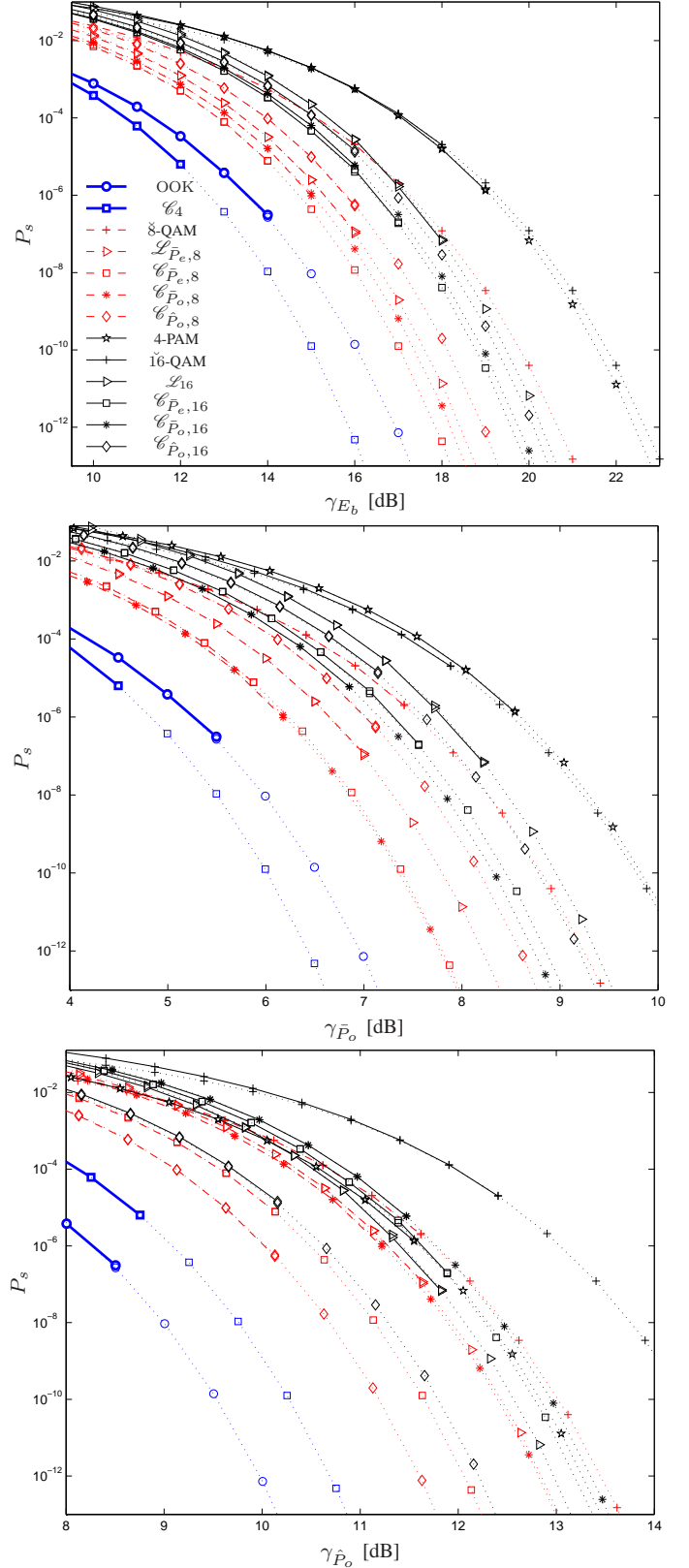


Fig. 6: Simulated (solid) and theoretical (dotted) SER for the modulation formats vs. γ_{E_b} (top), $\gamma_{\bar{P}_o}$ (middle), and $\gamma_{\hat{P}_o}$ (bottom) without coding.

zero-mean simplex constellation is $E_{s,\text{simplex}} = E_s/2 = E_b$, where E_s and E_b are the average symbol and bit energies, respectively, of the \mathcal{C}_4 constellation. Hence, the exact SER of \mathcal{C}_4 is

$$P_s = 1 - \int_{-\infty}^{\infty} \left[1 - Q\left(u + \sqrt{\frac{8E_b}{3N_0}}\right) \right]^3 \frac{e^{-u^2/2}}{\sqrt{2\pi}} du. \quad (24)$$

For higher-level modulation formats, the standard union bound found in [33, Eq. (4.81)] is used to approximate the theoretical SER. This union bound can be approximated as

$$P_s \approx \frac{2K}{M} Q\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right), \quad (25)$$

where K is the number of distinct signal pairs $(s_i(t), s_j(t))$ with $i < j$ for which $\int (s_i(t) - s_j(t))^2 dt = d_{\min}^2$. This approximation is tight at high SNR.

Fig. 6 (top) shows the simulated and theoretical SER of the studied modulation formats vs. electrical SNR defined as

$$\gamma_{E_b} = 10 \log_{10} \frac{E_b}{N_0} \quad [\text{dB}]. \quad (26)$$

As expected, all the modulation formats which are optimized for \bar{P}_e outperform the other formats at the same spectral efficiency. For spectral efficiency $\eta = 1$ bit/s/Hz (blue), \mathcal{C}_4 has a 0.86 dB average electrical power gain over OOK to achieve $P_s = 10^{-6}$. For $\eta = 1.5$ bit/s/Hz (red), $\mathcal{C}_{\bar{P}_e,8}$ has a 0.31 dB gain over $\mathcal{C}_{\bar{P}_o,8}$, 0.58 dB gain over $\mathcal{L}_{\bar{P}_e,8}$, 1.09 dB gain over $\mathcal{C}_{\hat{P}_o,8}$, and 2.55 dB gain over 8-QAM to achieve $P_s = 10^{-6}$. For $\eta = 2$ bit/s/Hz (black), $\mathcal{C}_{\bar{P}_e,16}$ has a 0.15 dB gain over $\mathcal{C}_{\bar{P}_o,16}$, 0.47 dB gain over $\mathcal{C}_{\hat{P}_o,16}$, 0.74 dB gain over \mathcal{L}_{16} , 2.65 dB gain over 4-PAM, and 2.80 dB gain over 16-QAM. The modulation formats optimized for \bar{P}_e and \bar{P}_o are close in performance to those optimized for \hat{P}_o and to the lattice-based formats.

In order to compare modulation formats in terms of their average optical power requirements, we define the average optical SNR as

$$\gamma_{\bar{P}_o} = 10 \log_{10} \frac{\bar{P}_o}{c\sqrt{R_b N_0}} \quad [\text{dB}] \quad (27)$$

in a similar fashion as in [2, Eq. (5)]. Using (8), (9), (26), $rc = 1$, and $E_s = E_b R_b T_s$, this expression can be written as

$$\gamma_{\bar{P}_o} = \frac{1}{2} \gamma_{E_b} + 10 \log_{10} \frac{\mathbb{E}[s_{I,1}]}{\sqrt{\mathbb{E}[\|s_I\|^2]}}, \quad (28)$$

where the first term depends on the regular (electrical) SNR and the second depends only on the constellation geometry. Fig. 6 (middle) shows the SER plotted vs. $\gamma_{\bar{P}_o}$. Quite obviously, the modulation formats optimized for \bar{P}_o perform better than the rest. For $\eta = 1$ bit/s/Hz, \mathcal{C}_4 has a 0.43 dB average optical power gain over OOK to achieve an SER of 10^{-6} . For $\eta = 1.5$ bit/s/Hz, $\mathcal{C}_{\bar{P}_e,8}$ has a 0.04 dB gain over $\mathcal{C}_{\bar{P}_o,8}$, 0.46 dB gain over $\mathcal{L}_{\bar{P}_e,8}$, 0.84 dB gain over $\mathcal{C}_{\hat{P}_o,8}$, and 1.35 dB gain over 8-QAM to achieve $P_s = 10^{-6}$. For $\eta = 2$ bit/s/Hz, $\mathcal{C}_{\bar{P}_e,16}$ has a 0.13 dB gain over $\mathcal{C}_{\bar{P}_o,16}$, 0.45 dB gain over $\mathcal{C}_{\hat{P}_o,16}$, 0.66 dB gain over \mathcal{L}_{16} , 1.36 dB gain over 16-QAM, and 1.44 dB gain over 4-PAM to achieve $P_s = 10^{-6}$. It can be noticed that

the 8-level modulation formats optimized for average electrical and optical power are very close in performance.

In a similar fashion as $\gamma_{\bar{P}_o}$, we define the peak optical SNR

$$\gamma_{\hat{P}_o} = 10 \log_{10} \frac{\hat{P}_o}{c\sqrt{R_b N_0}} \quad [\text{dB}] \quad (29)$$

in order to assess the different modulation formats under study in terms of their peak optical power requirements. Using (8), (18), (26), $rc = 1$, and $E_s = E_b R_b T_s$, this expression can be written as

$$\gamma_{\hat{P}_o} = \frac{1}{2} \gamma_{E_b} + 10 \log_{10} \frac{\max_i \left\{ s_{i,1} + \sqrt{2(s_{i,2}^2 + s_{i,3}^2)} \right\}}{\sqrt{\mathbb{E}[\|s_I\|^2]}}, \quad (30)$$

where the first term depends on the regular (electrical) SNR and the second depends only on the constellation geometry. Fig. 6 (bottom) shows the SER plotted vs. $\gamma_{\hat{P}_o}$. It is clear that the modulation formats optimized for peak optical power outperform the other formats. For $\eta = 1$ bit/s/Hz, OOK has a 0.82 dB peak optical power gain over \mathcal{C}_4 to achieve $P_s = 10^{-6}$. For $\eta = 1.5$ bit/s/Hz, $\mathcal{C}_{\bar{P}_e,8}$ has a 0.46 dB gain over $\mathcal{C}_{\bar{P}_o,8}$, 1.20 dB gain over $\mathcal{C}_{\hat{P}_o,8}$, 1.25 dB gain over $\mathcal{L}_{\bar{P}_e,8}$, and 1.72 dB gain over 8-QAM to achieve an SER of 10^{-6} . For $\eta = 2$ bit/s/Hz, $\mathcal{C}_{\bar{P}_e,16}$ has a 0.81 dB peak optical power gain over \mathcal{L}_{16} , 0.98 dB gain over 4-PAM, 1.00 dB gain over $\mathcal{C}_{\bar{P}_e,16}$, 1.16 dB gain over $\mathcal{C}_{\bar{P}_o,16}$, and 2.42 dB gain over 16-QAM to achieve $P_s = 10^{-6}$. Overall, modulation formats optimized for \hat{P}_o perform well in average-power limited systems. In [54], a similar conclusion was reached for the case of coherent optical systems.

B. Asymptotic Power Efficiency

At asymptotically high SNR, the performance difference in dB between the different modulation formats approaches constant values. We use OOK as a benchmark for power-efficiency on IM/DD channels as in [1], [2].

Fig. 7 (top) presents the average electrical power gain

$$\bar{P}_{e,\text{gain}} = 10 \log_{10} \frac{\bar{P}_{e,\text{OOK}}}{\bar{P}_e} \quad [\text{dB}] \quad (31)$$

of a modulation format in comparison to OOK in order to achieve the same error rate performance at asymptotically high SNR, where $\bar{P}_{e,\text{OOK}}$ and \bar{P}_e are the average electrical power of OOK and the modulation format under study, respectively. For spectral efficiency $\eta = 1$ bit/s/Hz, \mathcal{C}_4 has a 1.25 dB average electrical power gain over OOK. The overall trend at asymptotically high SNR is similar to the comparison of average electrical SNR γ_{E_b} in Sec. V-A to achieve $P_s = 10^{-6}$. For $\eta = 1.5$ bit/s/Hz, the 8-level modulation format optimized for \bar{P}_e , i.e., $\mathcal{C}_{\bar{P}_e,8}$, has the smallest average electrical power penalty of 0.67 dB compared to OOK. The asymptotic average electrical power gain of $\mathcal{C}_{\bar{P}_e,8}$ over other formats is larger than the gains at $P_s = 10^{-6}$ with the exception of having a similar gain when compared to $\mathcal{L}_{\bar{P}_e,8}$ and $\mathcal{C}_{\hat{P}_o,8}$ where both are lattice-based formats. For $\eta = 2$ bit/s/Hz, $\mathcal{C}_{\bar{P}_e,16}$ has a 2.42 dB penalty with respect to OOK at asymptotically high SNR. Its asymptotic electrical power gain is larger than the

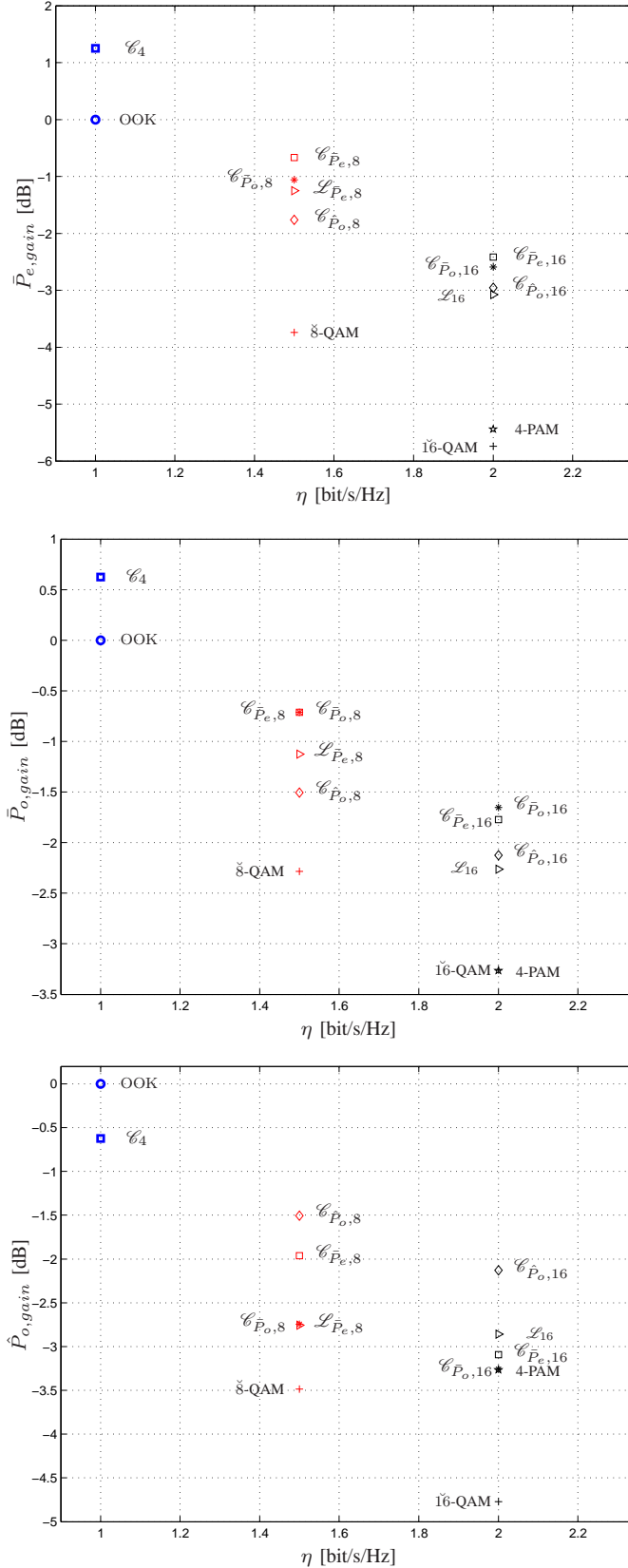


Fig. 7: Average electrical (top), average optical (middle), and peak optical (bottom) power gain of the various modulation formats vs. OOK without coding.

gains at $P_s = 10^{-6}$ except when compared to \mathcal{L}_{16} where the gain at asymptotically high SNR is smaller by 0.08 dB.

Fig. 7 (middle) shows the average optical power gain

$$\bar{P}_{o,gain} = 10 \log_{10} \frac{\bar{P}_{o,OOK}}{\bar{P}_o} \quad [\text{dB}] \quad (32)$$

of the modulation formats under study with respect to OOK at asymptotically high SNR, plotted versus their spectral efficiencies. The average optical power of the modulation format under study is denoted as \bar{P}_o . For $\eta = 1$ bit/s/Hz, \mathcal{C}_4 offers 0.62 dB average optical power gain over OOK. The asymptotic average optical power gain of \mathcal{C}_4 over OOK is larger than the gain at $P_s = 10^{-6}$ in Sec. V-A. For $\eta = 1.5$ bit/s/Hz, $\mathcal{C}_{\bar{P}_o,8}$ has a 0.71 dB average optical power penalty compared to OOK to achieve the same error rate at asymptotically high SNR. Furthermore, the asymptotic average optical power gains of $\mathcal{C}_{\bar{P}_o,8}$ are larger than the gains when compared with other formats to achieve $P_s = 10^{-6}$. Exceptions are when compared to $\mathcal{L}_{\bar{P}_e,8}$ and $\mathcal{C}_{\bar{P}_o,8}$, where the gains are smaller. Another observation is that the performance of $\mathcal{C}_{\bar{P}_o,8}$ becomes similar to that of $\mathcal{C}_{\bar{P}_e,8}$. For $\eta = 2$ bit/s/Hz, $\mathcal{C}_{\bar{P}_o,16}$ has a 1.65 dB penalty with respect to OOK at asymptotically high SNR. Compared to the average optical SNR performance gains at $P_s = 10^{-6}$, the asymptotic gains are larger except the gain over $\mathcal{C}_{\bar{P}_e,16}$ and \mathcal{L}_{16} .

If peak power is the limiting factor in a communication system, Fig. 7 (bottom) shows the performance of the various modulation formats with respect to their peak optical power gain with respect to OOK

$$\hat{P}_{o,gain} = 10 \log_{10} \frac{\hat{P}_{o,OOK}}{\hat{P}_o} \quad [\text{dB}], \quad (33)$$

where $\hat{P}_{o,OOK}$ and \hat{P}_o denote the peak optical power of OOK and the modulation format under study, respectively. For $\eta = 1$ bit/s/Hz, OOK has the best performance. It has a peak optical power gain of 0.62 dB over \mathcal{C}_4 . The asymptotic peak optical power gain of OOK compared to \mathcal{C}_4 is less than the gain in peak optical SNR to achieve $P_s = 10^{-6}$. In other words, the peak optical performance of \mathcal{C}_4 gets closer to OOK at asymptotically high SNR. For $\eta = 1.5$ bit/s/Hz, $\mathcal{C}_{\hat{P}_o,8}$ has the smallest penalty of 1.51 dB compared to OOK. The asymptotic peak optical power gain of $\mathcal{C}_{\hat{P}_o,8}$ over other formats is larger than the gain reported in Sec. V-A. The exception is the performance of $\mathcal{C}_{\hat{P}_o,8}$ when compared to $\mathcal{C}_{\bar{P}_e,8}$ and $\mathcal{L}_{\bar{P}_e,8}$, where the asymptotic peak optical power gain is similar to that at $P_s = 10^{-6}$. For $\eta = 2$ bit/s/Hz, $\mathcal{C}_{\hat{P}_o,16}$ outperforms the other modulation formats with a penalty of 2.13 dB with respect to OOK at asymptotically high SNR. Even though the asymptotic peak optical power gain of $\mathcal{C}_{\hat{P}_o,16}$ over most other formats is larger than that at $P_s = 10^{-6}$, it is smaller when $\mathcal{C}_{\hat{P}_o,16}$ is compared to \mathcal{L}_{16} , $\mathcal{C}_{\bar{P}_e,16}$, and $\mathcal{C}_{\bar{P}_o,16}$. So far, all the differences between the gains to achieve $P_s = 10^{-6}$ and the gains at asymptotically high SNR did not change the order of which formats have a better performance. However, 4-PAM at asymptotically high SNR is now worse than $\mathcal{C}_{\bar{P}_e,16}$ and $\mathcal{C}_{\bar{P}_o,16}$ in terms of peak optical power. Finally, the trend that can be observed in Fig. 7 (bottom) is that the gap in performance

between the modulation formats optimized for peak optical power and the rest of the formats under study gets larger with higher spectral efficiencies.

C. Mutual Information vs. SNR

Fig. 8 depicts the spectral efficiency where $R = I(\mathbf{x}; \mathbf{y})$ of the modulation formats vs. the different SNR measures, i.e., average electrical, average optical, and peak optical SNR. We study the same modulation formats as before, designed to minimize the uncoded SER, although they are now evaluated in a coding context. The mutual information $I(\mathbf{x}; \mathbf{y})$ for a fixed constellation and input distribution gives a lower bound on the rate in bits per symbol at which information can be sent with arbitrarily low probability of error, i.e., with the use of optimal coding. Throughout this paper, we consider uniform input distributions.

In Fig. 8 (top), the spectral efficiency is plotted vs. γ_{E_b} . The best format in terms of γ_{E_b} is \mathcal{C}_4 for $0 < \eta < 0.74$, $\mathcal{C}_{\bar{P}_o,8}$ for $0.74 < \eta < 0.99$, and $\mathcal{C}_{\bar{P}_o,16}$ for $\eta > 0.99$. It is surprising that the formats optimized for \bar{P}_o outperform the formats optimized for \bar{P}_e , when compared in terms of γ_{E_b} . This is different from the uncoded case in Sec. V-A-V-B. The 8-QAM and $\mathcal{L}_{\bar{P}_e,8}$ formats have a better performance than $\mathcal{C}_{\bar{P}_e,8}$. Other observations are that $\mathcal{C}_{\bar{P}_o,16}$ outperforms \mathcal{L}_{16} , and 4-PAM has the lowest performance.

In a similar fashion as above, Fig. 8 (middle) shows the spectral efficiency η plotted vs. the SNR $\gamma_{\bar{P}_o}$ for the same modulation formats. The best format is \mathcal{C}_4 for $0 < \eta < 0.83$, $\mathcal{C}_{\bar{P}_o,8}$ for $0.83 < \eta < 1.18$, and $\mathcal{C}_{\bar{P}_o,16}$ for $\eta > 1.18$. As opposed to the performance vs. γ_{E_b} , the gap between the formats optimized for average optical power and those for average electrical power is larger. It is interesting that the modulation formats optimized for average optical power perform better than the rest for systems which are limited by either average electrical or optical power.

Fig. 8 (bottom) presents the spectral efficiency vs. $\gamma_{\bar{P}_o}$. But here, the story is a bit different. The best format is OOK for $0 < \eta < 0.93$, and $\mathcal{C}_{\bar{P}_o,16}$ above that. An interesting observation is that 4-PAM performs better than $\mathcal{C}_{\bar{P}_o,16}$ for $0 < \eta < 0.56$. Furthermore, modulation formats optimized for peak power perform better in average-power limited systems than when using formats optimized for average power in peak-power limited systems.

D. Mutual Information in the Wideband Regime

It is apparent from the results in Sec. V-C that constellations optimized for a minimum-distance criterion (i.e., uncoded transmission at high SNR) do not necessarily perform well in terms of mutual information, particularly not in the wideband regime (low spectral efficiency). We define the zero-crossing $\nu(\Omega)$ of a constellation Ω in terms of an SNR measure as the minimum SNR for which $\eta > 0$. For the three SNR measures considered in this work (γ_{E_b} , $\gamma_{\bar{P}_o}$, or $\gamma_{\bar{P}_e}$), the zero-crossings will be denoted as $\nu_{E_b}(\Omega)$, $\nu_{\bar{P}_o}(\Omega)$, and $\nu_{\bar{P}_e}(\Omega)$, respectively. Fig. 8 shows that in terms of γ_{E_b} and $\gamma_{\bar{P}_o}$, \mathcal{C}_4 and OOK have the same lowest zero-crossing of 1.42 dB and -0.79 dB, respectively, among all studied formats. However, OOK

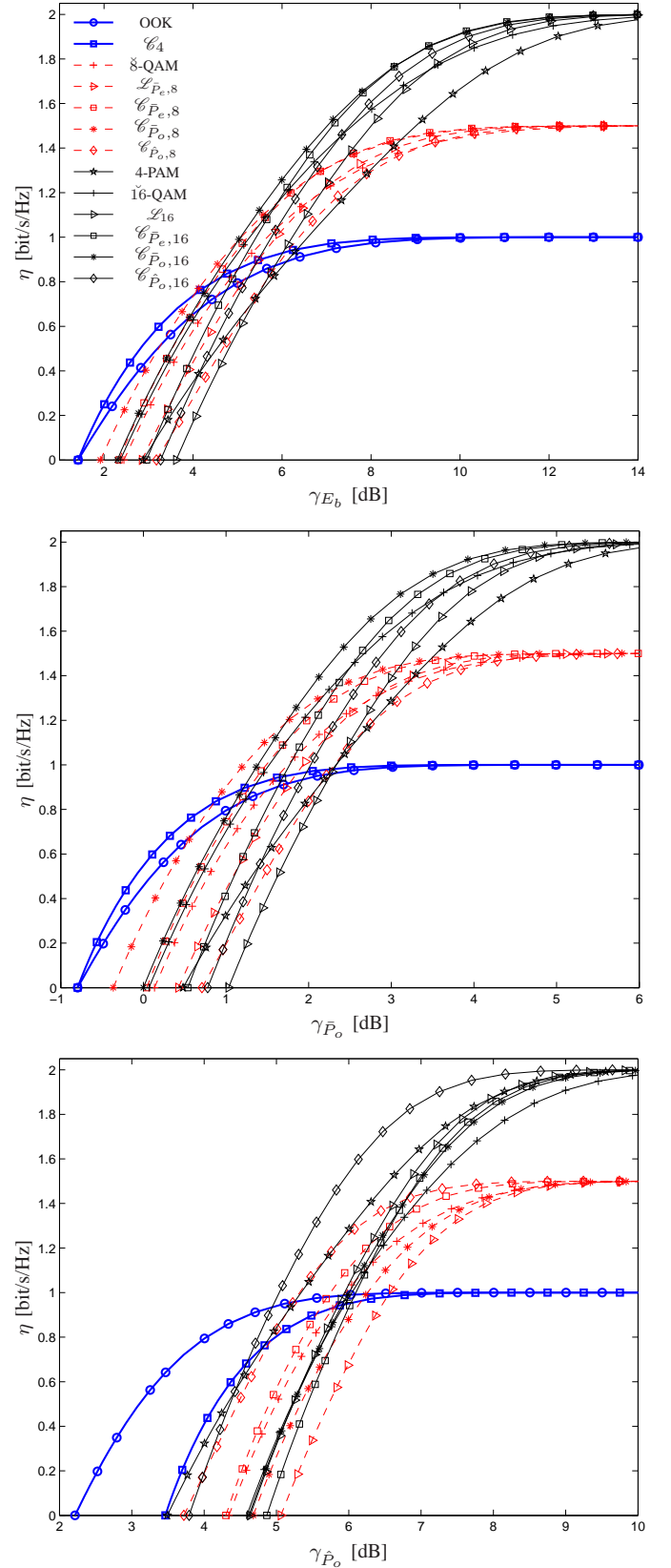


Fig. 8: Spectral efficiency vs. γ_{E_b} (top), $\gamma_{\bar{P}_o}$ (middle), and $\gamma_{\bar{P}_e}$ (bottom) in the presence of capacity-achieving codes.

has the lowest zero-crossing of 2.21 dB in terms of $\gamma_{\hat{P}_o}$. The zero-crossings will be explained analytically in this section. An interesting question is if there exist other constellations that perform better in the wideband regime than those found in Sec. IV. As we shall see in this section, the answer is yes.

Theorem 5: For any given $M \geq 2$ with $|\Omega| = M$ and a uniform input distribution,

$$\nu_{E_b}(\Omega) = 10 \log_{10} \left(\left(1 - \frac{\|\sum_{i=0}^{M-1} \mathbf{s}_i\|^2}{M \sum_{i=0}^{M-1} \|\mathbf{s}_i\|^2} \right)^{-1} \log_e 2 \right), \quad (34)$$

$$\nu_{\bar{P}_o}(\Omega) = 5 \log_{10} \left(\frac{(\sum_{i=0}^{M-1} s_{i,1})^2}{M \sum_{i=0}^{M-1} \|\mathbf{s}_i\|^2 - \|\sum_{i=0}^{M-1} \mathbf{s}_i\|^2} \log_e 2 \right), \quad (35)$$

$$\begin{aligned} \nu_{\bar{P}_o}(\Omega) &= 5 \log_{10} \left(\frac{\left(M \max_i \left\{ s_{i,1} + \sqrt{2(s_{i,2}^2 + s_{i,3}^2)} \right\} \right)^2}{M \sum_{i=0}^{M-1} \|\mathbf{s}_i\|^2 - \|\sum_{i=0}^{M-1} \mathbf{s}_i\|^2} \log_e 2 \right). \end{aligned} \quad (36)$$

Proof: According to [55, Th. 7], the zero-crossing

$$\nu_{E_b}(\Omega) = 10 \log_{10} \frac{1}{\alpha(\Omega)} \quad [\text{dB}], \quad (37)$$

where

$$\alpha(\Omega) = \left(1 - \frac{\|\mathbb{E}[\Omega]\|^2}{E_s} \right) \log_2 e, \quad (38)$$

and $\mathbb{E}[\Omega]$ is the mean of the constellation. Substituting (38) in (37) yields (34), whereas substituting (34) in (28) and (30) yields (35) and (36), respectively. ■

The problem of optimizing modulation formats which require the least SNR for reliable communications can be formulated as finding the constellation Ω which provide the minimum zero-crossing $\nu(\Omega)$. It can be represented mathematically as

$$\text{Minimize} \quad \nu(\Omega) \quad (39)$$

$$\text{Subject to} \quad |\Omega| = M \quad (40)$$

$$\Omega \subset \Upsilon, \quad (41)$$

for the three SNR measures considered in this work. Observe that no minimum-distance condition applies. Using [55, Th. 7], this optimization problem can be solved analytically.

Theorem 6: For any given $M \geq 2$ and $\Omega \subset \Upsilon$ with $|\Omega| = M$,

$$\nu_{E_b}(\Omega) \geq 10 \log_{10} \frac{M \log_e 2}{M-1} \quad [\text{dB}], \quad (42)$$

with equality if and only if $M-1$ constellation points are at the origin.

Proof: The numerator of (34) can be bounded using

$$\left\| \sum_{i=0}^{M-1} \mathbf{s}_i \right\|^2 = \sum_{i=0}^{M-1} \|\mathbf{s}_i\|^2 + 2 \sum_{i=0}^{M-2} \sum_{j=i+1}^{M-1} \langle \mathbf{s}_i, \mathbf{s}_j \rangle \quad (43)$$

$$\geq \sum_{i=0}^{M-1} \|\mathbf{s}_i\|^2, \quad (44)$$

where $\langle \cdot, \cdot \rangle$ for $\mathbf{s}_i, \mathbf{s}_j \in \Upsilon$, denotes the inner product, and (44) is due to the fact that

$$\langle \mathbf{s}_i, \mathbf{s}_j \rangle = \|\mathbf{s}_i\| \|\mathbf{s}_j\| \cos \theta \geq 0,$$

$\forall i, j = 0, \dots, M-1$, since from Th. 1, $\max \theta = \cos^{-1}(1/3) = 70.528^\circ \leq 90^\circ$. Equality holds if and only if $M-1$ points are located at the origin. Applying (44) in (34) completes the proof. ■

Several constellations fulfill Th. 6. If the nonzero constellation point has the coordinates $(\sqrt{ME_s}, 0, 0)$, we refer to this constellation as \mathcal{E}_M . At the same symbol rate, \mathcal{E}_M occupies a bandwidth of $W = R_s$, as OOK and M -PAM. However, if the nonzero constellation point is at the surface of Υ , this point has coordinates $(\sqrt{(2/3)ME_s}, 0, \sqrt{ME_s/3})$, and will be referred to as \mathcal{O}_M . The constellation \mathcal{O}_M belongs to the SCM family; therefore it occupies a bandwidth of $W = 2R_s$, which is twice the bandwidth of OOK and M -PAM at the same symbol rate. The reason for choosing the two extremes \mathcal{E}_M and \mathcal{O}_M is that both have the same \bar{P}_e ; however, \mathcal{O}_M has lower \bar{P}_o (see Fig. 2) and \mathcal{E}_M has lower bandwidth.

Theorem 7: For any given $M \geq 2$ and $\Omega \subset \Upsilon$ with $|\Omega| = M$,

$$\nu_{\bar{P}_o}(\Omega) \geq 5 \log_{10} \left(\frac{2}{3(M-1)} \log_e 2 \right) \quad [\text{dB}], \quad (45)$$

with equality if and only if $\Omega = \mathcal{O}_M$.

Proof: By using (35), $\nu_{\bar{P}_o}(\Omega)$ can be bounded by

$$\nu_{\bar{P}_o}(\Omega) \geq 5 \log_{10} \left(\frac{\sum_{i=0}^{M-1} s_{i,1}^2}{M \sum_{i=0}^{M-1} \|\mathbf{s}_i\|^2 - \|\sum_{i=0}^{M-1} \mathbf{s}_i\|^2} \log_e 2 \right) \quad (46)$$

$$\geq 5 \log_{10} \left(\frac{\sum_{i=0}^{M-1} s_{i,1}^2}{(M-1) \sum_{i=0}^{M-1} \|\mathbf{s}_i\|^2} \log_e 2 \right) \quad (47)$$

$$\geq 5 \log_{10} \left(\frac{2}{3(M-1)} \log_e 2 \right). \quad (48)$$

The inequality in (46) follows from $s_{i,1}, s_{j,1} \geq 0, \forall i, j = 0, 1, \dots, M-1$, whereas (47) follows from (44) and (48) follows from the definition of the admissible region in (17). This completes the proof of (45).

To prove that (45) is tight if and only if $\Omega = \mathcal{O}_M$, we observe that (46) and (47) are tight if and only if $M-1$ points are at the origin, while (48) is tight if and only if all M points are located on the boundary of the cone. ■

So far, we were able to derive optimal constellations in terms of ν_{E_b} and $\nu_{\bar{P}_o}$. For the third SNR measure $\nu_{\hat{P}_o}$, however, we resort to a conjecture only, for which we have solid numerical support but no proof.

Conjecture 8: For any $\Omega \subset \Upsilon$,

$$\nu_{\hat{P}_o} \geq 5 \log_{10}(4 \log_e 2), \quad (49)$$

with equality if and only if $\Omega = \mathcal{E}_2$ (OOK).

Fig. 9 (top) shows the spectral efficiency in the low γ_{E_b} regime for \mathcal{E}_M and \mathcal{O}_M , where $M = 4, 8$, and 16. We also include some of the previously optimized constellations for comparison. Even though \mathcal{E}_M and \mathcal{O}_M , for a given M , have

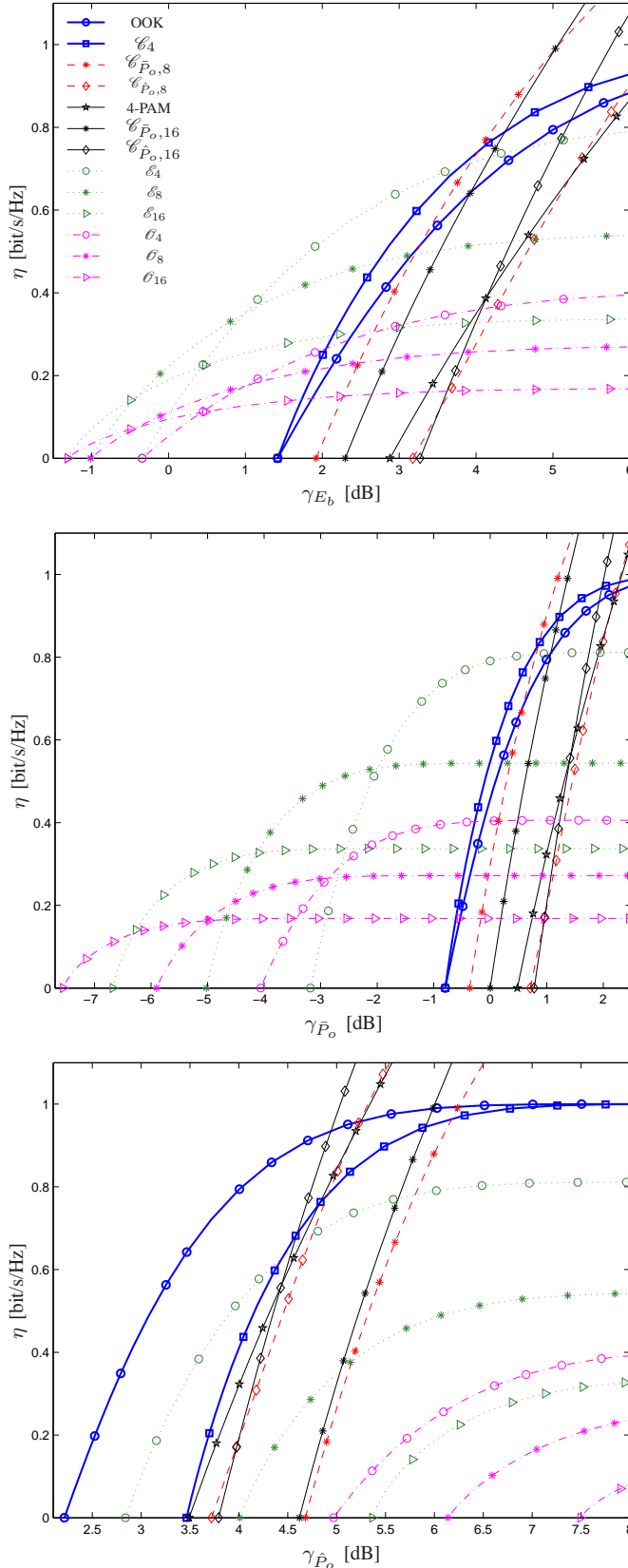


Fig. 9: Spectral efficiency of the optimal constellations \mathcal{E}_M and \mathcal{O}_M in the wideband regime. Some of the previously optimized constellations are included for comparison.

the same zero-crossing ν_{E_b} , \mathcal{E}_M performs better at all nonzero spectral efficiencies. Indeed, \mathcal{E}_{16} , \mathcal{E}_8 , and \mathcal{E}_4 are the best of all studied constellations for $\eta < 0.70$. As $M \rightarrow \infty$, \mathcal{E}_M and \mathcal{O}_M approach the coherent (non-IM/DD) Shannon limit, i.e., $\nu_{E_b}(\mathcal{E}_M) = \nu_{E_b}(\mathcal{O}_M) \rightarrow -1.59$ dB.

Fig. 9 (middle) depicts the spectral efficiency in the low $\gamma_{\bar{P}_o}$ regime. For $M = 4, 8$, and 16 , the formats \mathcal{O}_M have lower $\nu_{\bar{P}_o}(\Omega)$ than \mathcal{E}_M ; however, the latter performs better than all other studied formats for $0.13 < \eta < 0.80$. Using (45), $\nu_{\bar{P}_o}(\Omega)$ approaches $-\infty$ as $M \rightarrow \infty$ faster for $\Omega = \mathcal{O}_M$ than \mathcal{E}_M .

Finally, we show the spectral efficiency in the low $\gamma_{\bar{P}_o}$ regime in Fig. 9 (bottom). Surprisingly, OOK has the best performance up to $\eta = 0.93$, and then $\mathcal{E}_{\bar{P}_o,16}$ outperforms the rest. On the other hand, the modulation formats which minimize $\nu_{E_b}(\Omega)$ or $\nu_{\bar{P}_o}(\Omega)$ do not have the best performance as before. The trend is that their performance is poorer with increased number of levels. Also, \mathcal{O}_M is worse than \mathcal{E}_M for a given M , since the former has a higher peak power requirement due to the nonzero point on the surface of Υ .

It may seem somewhat unexpected that a constellation with $M - 1$ points at the origin would perform better at low SNR than all other constellations. Such a constellation is equivalent to a binary constellation with a higher probability for the zero point. Similar constellations were found to achieve capacity in [42] and [56], albeit for different channels. A similar constellation was studied in [57], where a signal set for coherent AWGN channels consisting of two antipodal signals and $M - 2$ signals at the origin was used to disprove the strong simplex conjecture, according to which the regular simplex signal set would minimize the uncoded probability of error under an average energy constraint.

VI. CONCLUSIONS

By using the minimum distance as a modulation design criterion for uncoded systems, we were able to numerically optimize 4-, 8-, and 16-level single-subcarrier IM/DD modulation formats for systems which are limited by average electrical, average optical, and peak power. For $M = 4$, the most power-efficient modulation in terms of average electrical, optical, and peak power has a tetrahedral structure. This constellation is also a subset of all the obtained higher-level constellations. As for the 8- and 16-level constellations, power-efficient schemes are obtained by not confining the set of constellation points to a regular structure such as that of a lattice. However, this comes at the price of losing the geometric regularity, which increases the modulator and demodulator complexity. Our comparisons show that the penalty gap for using modulation formats optimized for \bar{P}_o in systems which are \bar{P}_e or \bar{P}_o limited is much less than the penalty gap when using formats optimized for \bar{P}_e or \bar{P}_o in systems which are limited by \bar{P}_o . Therefore, modulation formats optimized for \bar{P}_o should be preferred in applications with mixed power requirements. The overall gain of the obtained formats over the previously best known formats is between 0.6 dB and 3 dB at asymptotically high SNR. We conjecture that the new obtained modulation formats are optimal for their size and optimization criteria over uncoded single-subcarrier IM/DD channels.

On the other hand, when capacity-achieving error-correcting codes are deployed, the best modulation formats in the wide-band regime (small η) have only one nonzero constellation point. This is confirmed analytically and numerically. At higher η , modulation formats optimized for \bar{P}_o are able to achieve higher reliable transmission rates compared to other formats in systems which are limited by average electrical or optical power. OOK and the 16-level modulation format optimized for peak power offer the best performance in peak-power limited systems. The overall gain of the obtained formats over previously best known formats ranges between 0.3 dB and 1 dB.

APPENDIX A OBTAINED CONSTELLATIONS

This appendix lists the coordinates of the numerically optimized constellations in Sec. IV-A. Whenever possible, numerical values have been replaced with the corresponding exact values. Constellations are normalized to unit d_{\min} .

$$\mathcal{C}_4 = \mathcal{C}_{\bar{P}_e,4} = \mathcal{C}_{\bar{P}_o,4} = \mathcal{C}_{\hat{P}_o,4} = \mathcal{L}_{\bar{P}_e,4} = \mathcal{L}_{\bar{P}_o,4} = \mathcal{L}_{\hat{P}_o,4} = \{(0, 0, 0), (\sqrt{2/3}, 0, 1/\sqrt{3}), (\sqrt{2/3}, \pm 1/2, -\sqrt{3}/6)\}.$$

$$\mathcal{C}_{\bar{P}_e,8} = \mathcal{C}_4 \cup \{((5/3)\sqrt{2/3}, 0, -5/(3\sqrt{3})), ((5/3)\sqrt{2/3}, \pm 5/6, 5/(6\sqrt{3})), (2\sqrt{2/3}, 0, 0)\}.$$

$$\mathcal{C}_{\bar{P}_o,8} = \mathcal{C}_4 \cup \{((5/3)\sqrt{2/3}, 0, -5/(3\sqrt{3})), ((5/3)\sqrt{2/3}, \pm 5/6, 5/(6\sqrt{3})), (1.6293, -0.9236, -0.6886)\}.$$

$$\mathcal{C}_{\hat{P}_o,8} = \mathcal{L}_{\bar{P}_o,8} = \mathcal{C}_4 \cup \{(2\sqrt{2/3}, 0, -1/\sqrt{3}), (2\sqrt{2/3}, \pm 1/2, \sqrt{3}/6), (\sqrt{6}, 0, 0)\}.$$

$$\mathcal{L}_{\bar{P}_e,8} = \mathcal{L}_{\bar{P}_o,8} = \mathcal{C}_4 \cup \{(2\sqrt{2/3}, \pm 1/2, \sqrt{3}/6), (2\sqrt{2/3}, 0, -1/\sqrt{3}), (2\sqrt{2/3}, 1, -1/\sqrt{3})\}.$$

$$\begin{aligned} \mathcal{C}_{\bar{P}_e,16} = \mathcal{C}_4 \cup \{ & (1.3608, 5/6, 5/(6\sqrt{3})), \\ & (1.3608, 0, -0.9623), (1.4628, -0.7513, 0.7110), \\ & (1.6024, -1.1134, -0.2106), (1.6293, 0.1346, 1.1442), \\ & (1.6293, 0.9236, -0.6887), (2\sqrt{2/3}, 0, 0), \\ & (1.9336, -0.8075, -1.1032), (2.0380, 1.4396, 0.0642), \\ & (2.3097, 0.5202, 0.5210), (2.3097, 0.1911, -0.7110), \\ & (2.3499, -0.6462, 0.2616)\}. \end{aligned}$$

$$\begin{aligned} \mathcal{C}_{\bar{P}_o,16} = \mathcal{C}_{\bar{P}_e,8} \cup \{ & (1.6293, -0.1345, 1.1442), \\ & (1.6293, 1.0582, -0.4556), (2\sqrt{2/3}, 0, 0), \\ & (2.0380, 0.6643, -1.2789), (2.0380, -1.4396, 0.0642), \\ & (2.0380, 0.7754, 1.2147), (2.1187, 1.4645, 0.3160), \\ & (2.1187, -1.0059, 1.1103)\}. \end{aligned}$$

$$\begin{aligned} \mathcal{C}_{\hat{P}_o,16} = \mathcal{C}_4 \cup \{ & (1.6279, 0.8995, -0.7184), \\ & (1.6279, -0.4977, 1.0379), \\ & (1.6270, -0.9003, -0.7162), (1.6300, 0.5022, 1.0374), \\ & (1.6310, -0.0010, -1.1533), \\ & (1.6313, -1.1242, 0.2584), (1.6328, 1.1259, 0.2557), \\ & (2\sqrt{2/3}, 0, 0), (2.4495, 0, 1/\sqrt{3}), \\ & (2.4495, \pm 1/2, -\sqrt{3}/6), (3.2660, 0, 0)\}. \end{aligned}$$

$$\begin{aligned} \mathcal{L}_{16} = \mathcal{L}_{\bar{P}_e,16} = \mathcal{L}_{\bar{P}_o,16} = \mathcal{L}_{\hat{P}_o,16} = \mathcal{L}_{\bar{P}_e,8} \cup \{ & 2\sqrt{2/3}, 0, (2/3)\sqrt{3}), \\ & (2\sqrt{2/3}, -1, -\sqrt{3}/3), (\sqrt{6}, 0, 0), \\ & (\sqrt{6}, -1/2, \sqrt{3}/2), (\sqrt{6}, \pm 1/2, -\sqrt{3}/2), \\ & (\sqrt{6}, \pm 1, 0)\}. \end{aligned}$$

ACKNOWLEDGMENT

The authors would like to acknowledge LINDO Systems for the free license to use their numerical optimization software. We would also like to acknowledge Dr. Giuseppe Durisi and Dr. Alex Alvarado for the interesting discussions about mutual information, Tilak Rajesh Lakshmana for the nice discussions about optimization techniques, and Rajet Krishnan for his comments about the paper structure.

REFERENCES

- [1] J. R. Barry, *Wireless Infrared Communications*. Norwell, MA, USA: Kluwer Academic Publishers, 1994.
- [2] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, 1997.
- [3] S. Hranilovic, *Wireless Optical Communication Systems*. New York: Springer, 2005.
- [4] S. Randel, F. Breyer, and S. C. J. Lee, "High-speed transmission over multimode optical fibers," in *Proc. Optical Fiber Communication Conference*, 2008, p. OWR2.
- [5] D. Molin, G. Kuyt, M. Bigot-Astruc, and P. Sillard, "Recent advances in MMF technology for data networks," in *Proc. Optical Fiber Communication Conference*, 2011, p. OWJ6.
- [6] S. Hranilovic and F. R. Kschischang, "Capacity bounds for power- and band-limited optical intensity channels corrupted by Gaussian noise," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 784–795, 2004.
- [7] —, "Optical intensity-modulated direct detection channels: Signal space and lattice codes," *IEEE Transactions on Information Theory*, vol. 49, no. 6, pp. 1385–1399, 2003.
- [8] A. A. Farid and S. Hranilovic, "Capacity bounds for wireless optical intensity channels with Gaussian noise," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6066–6077, 2010.
- [9] A. Lapidot, S. M. Moser, and M. A. Wigger, "On the capacity of free-space optical intensity channels," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.
- [10] J. E. Cunningham, D. Beckman, X. Zheng, D. Huang, T. Sze, and A. V. Krishnamoorthy, "PAM-4 signaling over VCSELs with 0.13 μ m CMOS chip technology," *Optics Express*, vol. 14, no. 25, pp. 12 028–12 038, 2006.

- [11] K. Szczerba, P. Westbergh, J. Karout, J. Gustavsson, Å. Haglund, M. Karlsson, P. Andrekson, E. Agrell, and A. Larsson, "30 Gbps 4-PAM transmission over 200 m of MMF using an 850 nm VCSEL," *Optics Express*, vol. 19, no. 26, pp. B203–B208, Dec. 2011.
- [12] S. Walklin and J. Conradi, "Multilevel signaling for increasing the reach of 10 Gb/s lightwave systems," *Journal of Lightwave Technology*, vol. 17, no. 11, pp. 2235–2248, 1999.
- [13] S. Hranilovic, "On the design of bandwidth efficient signalling for indoor wireless optical channels," *International Journal of Communication Systems*, vol. 18, no. 3, pp. 205–228, 2005.
- [14] A. O. J. Wiberg, B.-E. Olsson, and P. A. Andrekson, "Single cycle subcarrier modulation," in *Proc. Optical Fiber Communication Conference*, 2009, p. OTuE1.
- [15] B.-E. Olsson and A. Alping, "Electro-optical subcarrier modulation transmitter for 100 GbE DWDM transport," in *Proc. Asia Optical Fiber Communication and Optoelectronic Exposition and Conference*, 2008, p. SaF3.
- [16] B.-E. Olsson and M. Sköld, "QPSK transmitter based on optical amplitude modulation of electrically generated QPSK signal," in *Proc. Asia Optical Fiber Communication and Optoelectronic Exposition and Conference*, 2008, p. SaA3.
- [17] J. Armstrong, "OFDM for optical communications," *Journal of Lightwave Technology*, vol. 27, no. 3, pp. 189–204, Feb. 2009.
- [18] S. C. J. Lee, F. Breyer, S. Randel, H. P. A. van den Boom, and A. M. J. Koonen, "High-speed transmission over multimode fiber using discrete multitone modulation," *Journal of Optical Networking*, vol. 7, no. 2, pp. 183–196, Feb. 2008.
- [19] J. B. Carruthers and J. M. Kahn, "Multiple-subcarrier modulation for nondirected wireless infrared communication," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 3, pp. 538–546, Apr. 1996.
- [20] S. Randel, F. Breyer, S. C. J. Lee, and J. W. Walewski, "Advanced modulation schemes for short-range optical communications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1280–1289, Sep./Oct. 2010.
- [21] W. Kang and S. Hranilovic, "Optical power reduction for multiple-subcarrier modulated indoor wireless optical channels," in *Proc. IEEE International Conference on Communications*, 2006, pp. 2743–2748.
- [22] R. You and J. M. Kahn, "Average power reduction techniques for multiple-subcarrier intensity-modulated optical signals," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2164–2171, 2001.
- [23] S. Hranilovic and D. A. Johns, "A multilevel modulation scheme for high-speed wireless infrared communications," in *Proc. IEEE International Symposium on Circuits and Systems*, 1999, pp. 338–341.
- [24] D.-S. Shiu and J. M. Kahn, "Shaping and nonequiprobable signaling for intensity-modulated signals," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2661–2668, 1999.
- [25] G. D. Forney, Jr., "Coset codes—Part I: Introduction and geometrical classification," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 1123–1151, 1988.
- [26] G. D. Forney, Jr. and L.-F. Wei, "Multidimensional constellations—Part I. Introduction, figures of merit, and generalized cross constellations," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 877–892, 1989.
- [27] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. New York: Springer-Verlag, 1999.
- [28] A. R. Calderbank and L. H. Ozarow, "Nonequiprobable signaling on the Gaussian channel," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 726–740, 1990.
- [29] W. Mao and J. M. Kahn, "Lattice codes for amplified direct-detection optical systems," *IEEE Transactions on Communications*, vol. 56, no. 7, pp. 1137–1145, 2008.
- [30] J. Karout, E. Agrell, and M. Karlsson, "Power efficient subcarrier modulation for intensity modulated channels," *Optics Express*, vol. 18, no. 17, pp. 17913–17921, Aug. 2010.
- [31] K. Szczerba, J. Karout, P. Westbergh, E. Agrell, M. Karlsson, P. Andrekson, and A. Larsson, "Experimental comparison of modulation formats in IM/DD links," *Optics Express*, vol. 19, no. 10, pp. 9881–9889, May 2011.
- [32] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, Feb. 2010.
- [33] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques: Signal Design and Detection*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [34] L. P. Chen and K. Y. Lau, "Regime where zero-bias is the low-power solution for digitally modulated laser diodes," *IEEE Photonics Technology Letters*, vol. 8, no. 2, pp. 185–187, 1996.
- [35] C. Cox and W. S. C. Chang, "Figures of merit and performance analysis of photonic microwave links," in *RF Photonic Technology in Optical Fiber Links*, W. S. C. Chang, Ed. Cambridge University Press, 2002, ch. 1, pp. 1–33.
- [36] B. Inan, S. C. J. Lee, S. Randel, I. Neokosmidis, A. M. J. Koonen, and J. W. Walewski, "Impact of LED nonlinearity on discrete multitone modulation," *Journal of Optical Communications and Networking*, vol. 1, no. 5, pp. 439–451, Oct. 2009.
- [37] K.-P. Ho, *Phase-Modulated Optical Communication Systems*. New York: Springer, 2005.
- [38] P. Westbergh, J. S. Gustavsson, Å. Haglund, A. Larsson, F. Hopfer, G. Fiol, D. Bimberg, and A. Joel, "32 Gbit/s multimode fibre transmission using high-speed, low current density 850 nm VCSEL," *Electronics Letters*, vol. 45, no. 7, pp. 366–368, 2009.
- [39] L. A. Coldren and E. R. Hegblom, "Fundamental issues in VCSEL design," in *Vertical-Cavity Surface-Emitting Lasers: Design, Fabrication, Characterization, and Applications*, C. W. Wilmsen, H. Temkin, and L. A. Coldren, Eds. Cambridge University Press, 1999, ch. 2, pp. 32–67.
- [40] G. P. Agrawal, *Lightwave Technology: Telecommunication Systems*. New Jersey: John Wiley & Sons, Inc., 2005.
- [41] B. Goebel, R.-J. Essiambre, G. Kramer, P. J. Winzer, and N. Hanik, "Calculation of mutual information for partially coherent Gaussian channels with applications to fiber optics," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 5720–5736, Sep. 2011.
- [42] K.-P. Ho, "Exact evaluation of the capacity for intensity-modulated direct-detection channels with optical amplifier noises," *IEEE Photonics Technology Letters*, vol. 17, no. 4, pp. 858–860, Apr. 2005.
- [43] M. S. Moreolo, R. Muñoz, and G. Junyent, "Novel power efficient optical OFDM based on Hartley transform for intensity-modulated direct-detection systems," *Journal of Lightwave Technology*, vol. 28, no. 5, pp. 798–805, Mar. 2010.
- [44] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New Jersey: John Wiley & Sons, Inc., 2006.
- [45] G. Foschini, R. Gitlin, and S. Weinstein, "Optimization of two-dimensional signal constellations in the presence of Gaussian noise," *IEEE Transactions on Communications*, vol. COM-22, no. 1, pp. 28–38, Jan. 1974.
- [46] J.-E. Porath and T. Aulin, "Design of multidimensional signal constellations," *IEE Proceedings - Communications*, vol. 150, no. 5, pp. 317–323, Oct. 2003.
- [47] N. J. A. Sloane, R. H. Hardin, T. D. S. Duff, and J. H. Conway, "Minimal-energy clusters of hard spheres," *Discrete and Computational Geometry*, vol. 14, no. 3, pp. 237–259, 1995.
- [48] R. L. Graham and N. J. A. Sloane, "Penny-packing and two-dimensional codes," *Discrete and Computational Geometry*, vol. 5, no. 1, pp. 1–11, 1990.
- [49] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *Journal of Lightwave Technology*, vol. 27, no. 22, pp. 5115–5126, 2009.
- [50] J. R. Pierce, "Comparison of three-phase modulation with two-phase and four-phase modulation," *IEEE Transactions on Communications*, vol. COM-28, no. 7, pp. 1098–1099, 1980.
- [51] N. Ekanayake and T. T. Tjhung, "On ternary phase-shift keyed signaling," *IEEE Transactions on Information Theory*, vol. IT-28, no. 4, pp. 658–660, Jul. 1982.
- [52] M. C. Gursoy, "Error rate analysis for peaky signaling over fading channels," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2546–2550, Sep. 2009.
- [53] J. Karout, "Modulation Optimization for Noncoherent Optical Systems," Licentiate Thesis (R010/2011), Chalmers University of Technology, Gothenburg, Sweden, 2011.
- [54] M. Karlsson and E. Agrell, "Power-efficient modulation schemes," in *Impact of Nonlinearities on Fiber Optic Communications*, S. Kumar, Ed. Springer, 2011, ch. 5, pp. 219–252.
- [55] E. Agrell and A. Alvarado, "Optimal signal sets and binary labelings for BICM at low SNR," *IEEE Transactions on Information Theory*, to appear, Jun. 2011, arXiv:1001.4548v2.
- [56] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Information and Control*, vol. 18, pp. 203–219, 1971.
- [57] M. Steiner, "The strong simplex conjecture is false," *IEEE Transactions on Information Theory*, vol. 40, no. 3, pp. 721–731, May 1994.